Internet Worm and Virus Protection for Very High-Speed Networks

John W. Lockwood
Professor of Computer Science and Engineering

lockwood@arl.wustl.edu
http://www.arl.wustl.edu/~lockwood

Research Sponsor:

http://www.globalvelocity.info/

Internet Worms and Viruses

- The problem with worms and virus attacks
  - Annoyance to users
  - Costly to businesses (lost productivity)
  - Security threat to government (compromised data)

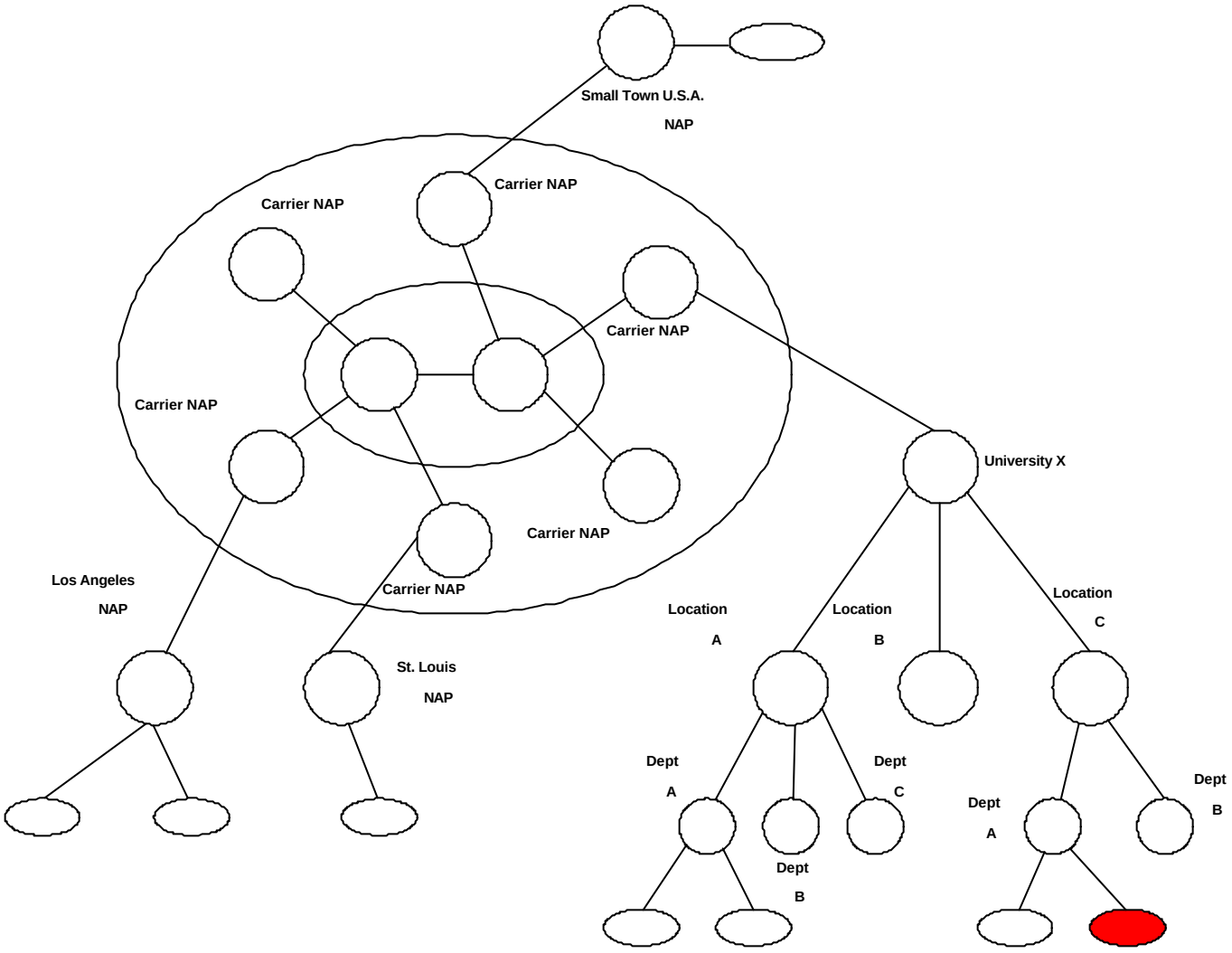
- Recent Attacks
  - Nimda, Code Red, Slammer
  - MSBlast
    - Infected over 350,000 hosts in Aug. 16, 2003
  - SoBigF
    - Infected 1 million users in first 24 hours
    - Infected > 200 million in the first week
    - Caused an estimated $1 billion in damages to repair.

- Detectable by a Signature in Content
  - Pattern of bytes
  - Regular Expression
  - Morphable pattern

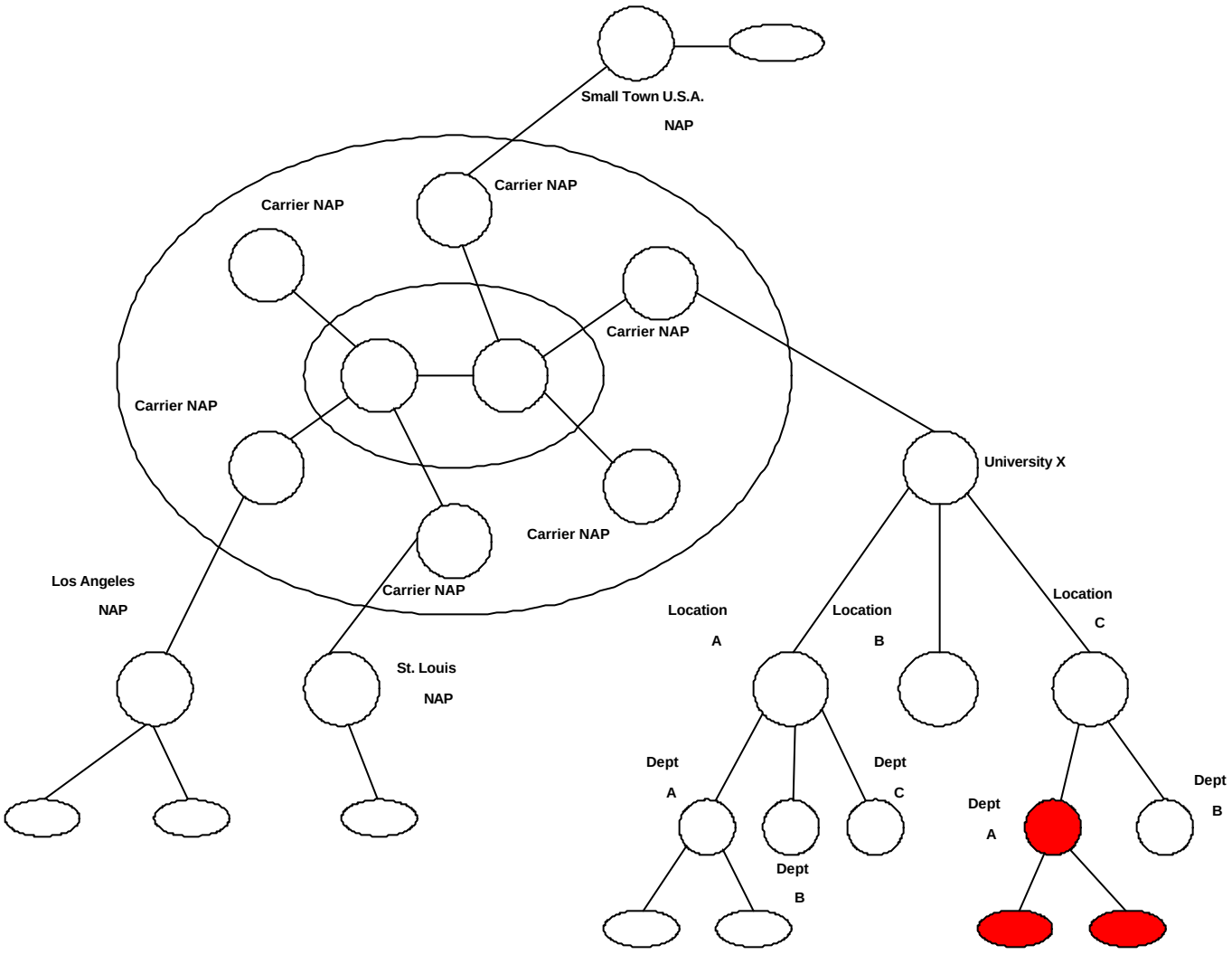
Challenges to Stopping Worm and Virus Attacks

- End-systems difficult to maintain
  - Operating systems become outdated
  - Users introduce new machines on network
- Internet contains several types of traffic
  - Web, file transfers, telnet
  - Data may appear anywhere in the packet
- Networks process High Speed Data
  - Multi Gigabit/second data transmission rates now commonplace in campus, corporate, and backbone networks
  - Peer-to-Peer protocols dominate current and future traffic
  - Need Real-time gathering
    - No latency can be tolerated

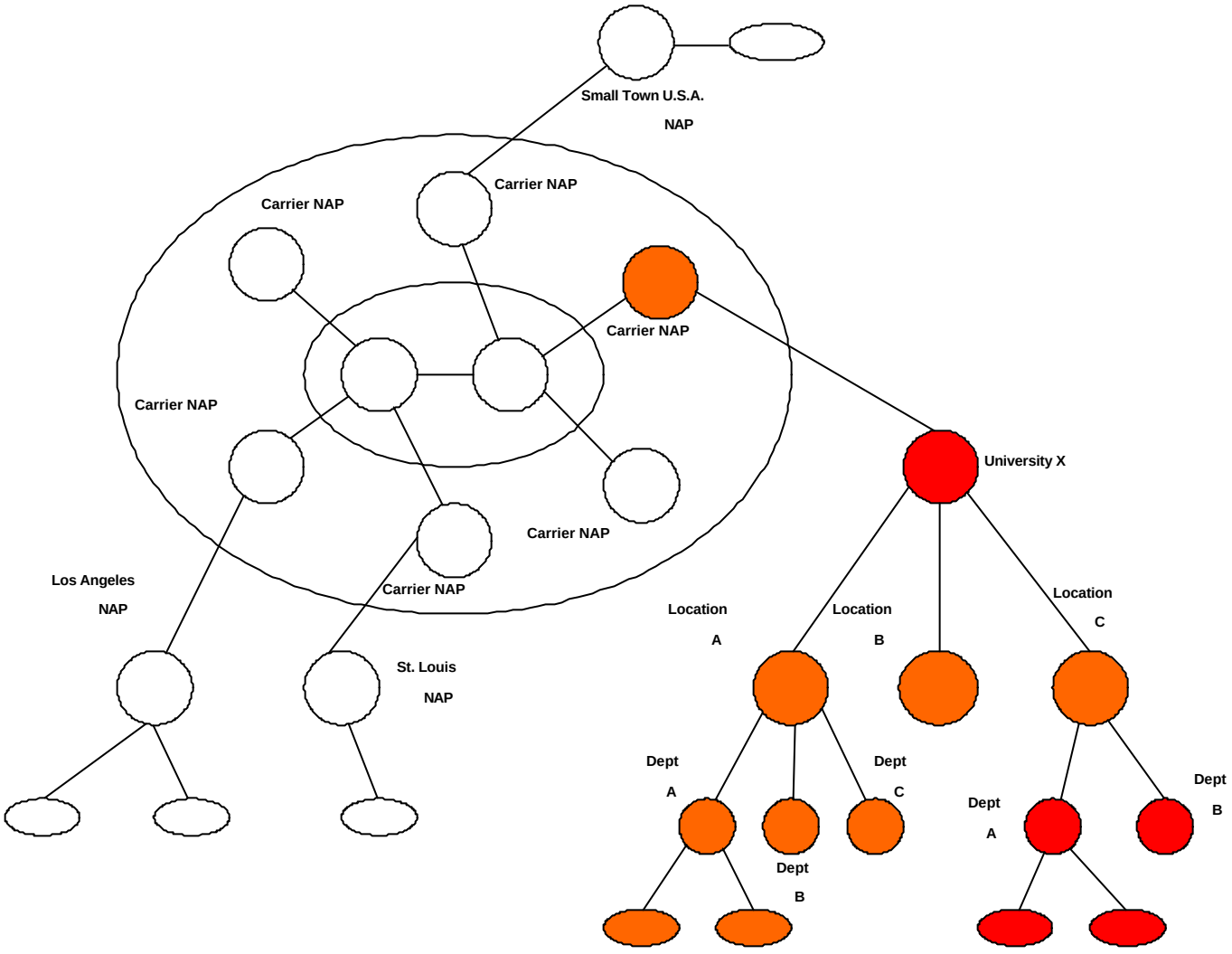
Virus/Worm/Data Spread in Unprotected Networks



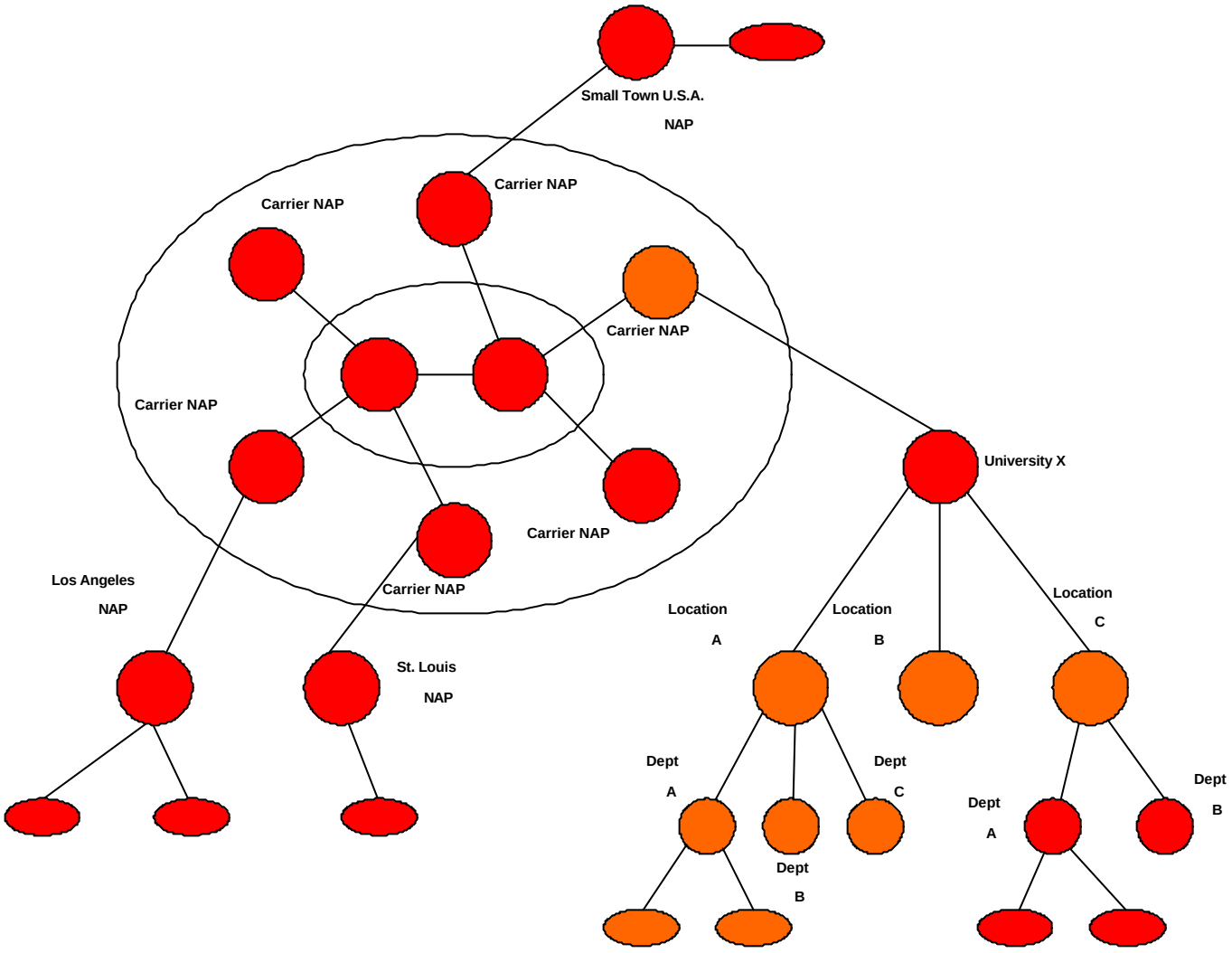
Virus/Worm/Data Spread in Unprotected Networks



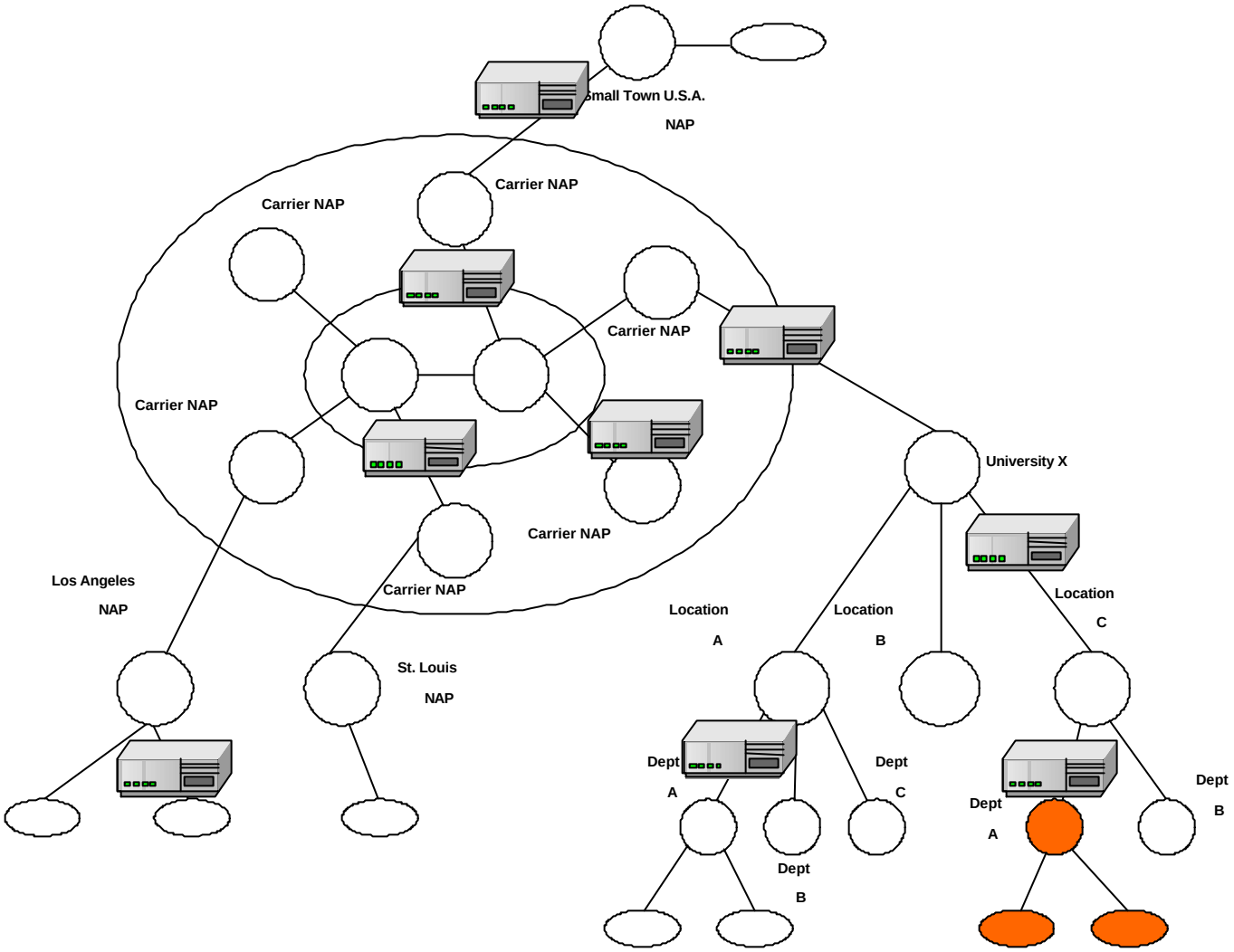
Virus/Worm/Data Spread in Unprotected Networks



Virus/Worm/Data Spread in Unprotected Networks



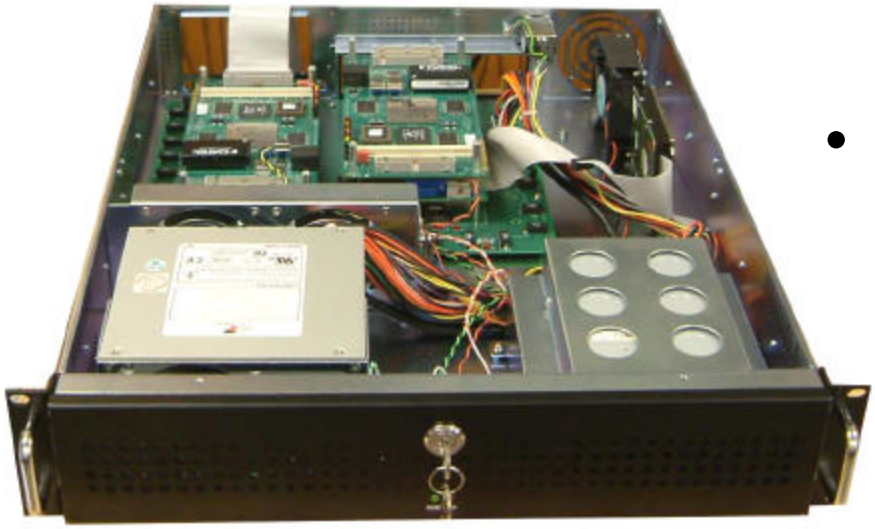
Virus/Worm/Data Containment in Protected Networks



Content Scanning and Protection Device

Washington University in St. Louis

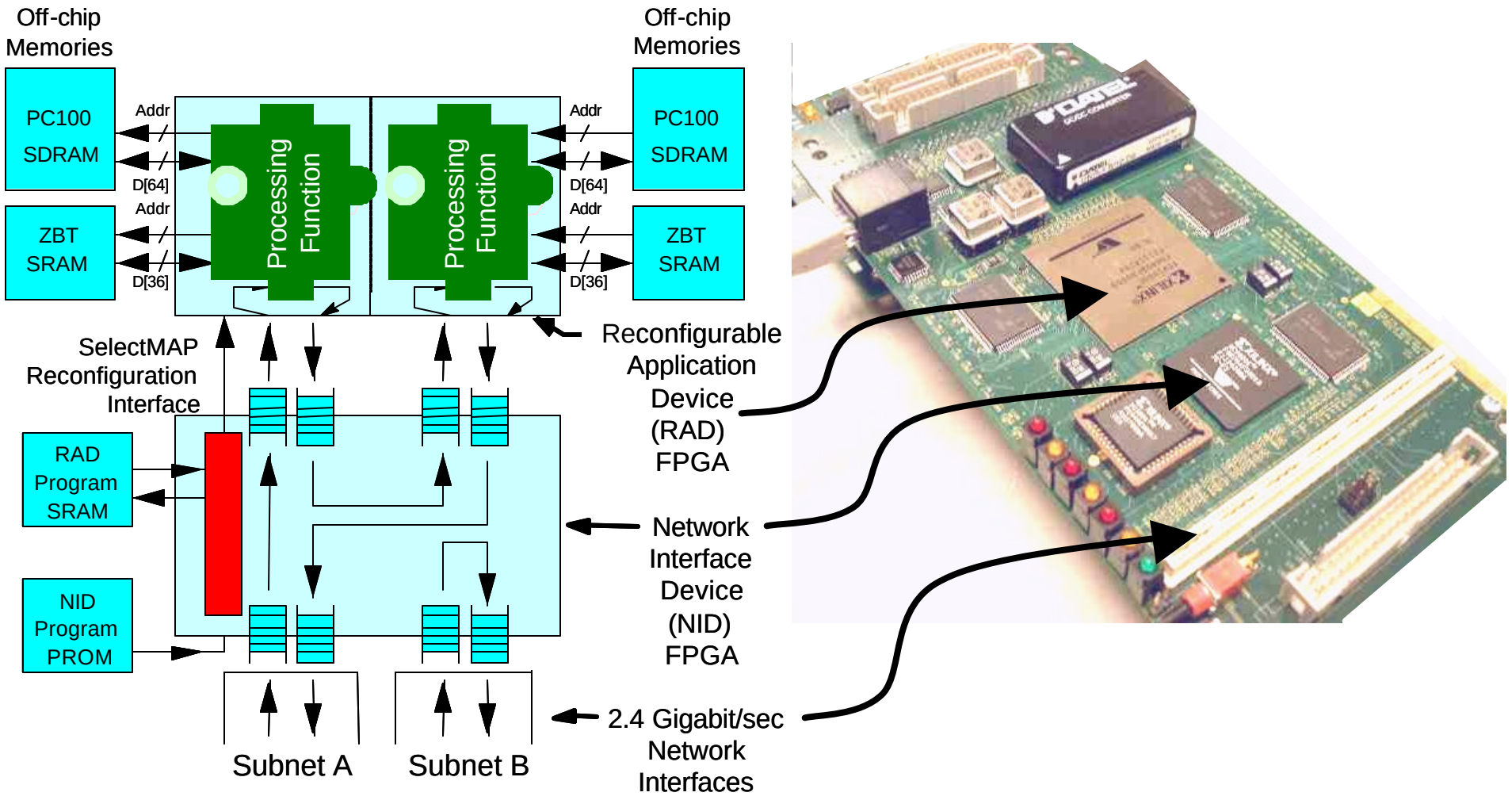
Content Scanning Technology



- Fiber optic Line Cards
  - Gigabit Ethernet
  - ATM OC-3 to OC-48
- Reconfigurable Hardware
  - Uses Field Programmable Port Extender (FPX) Platform
  - Protocol processing and content scanning performed in hardware
  - Reconfigurable over the network
- Chassis / Motherboard
  - Allows Modules to Stack

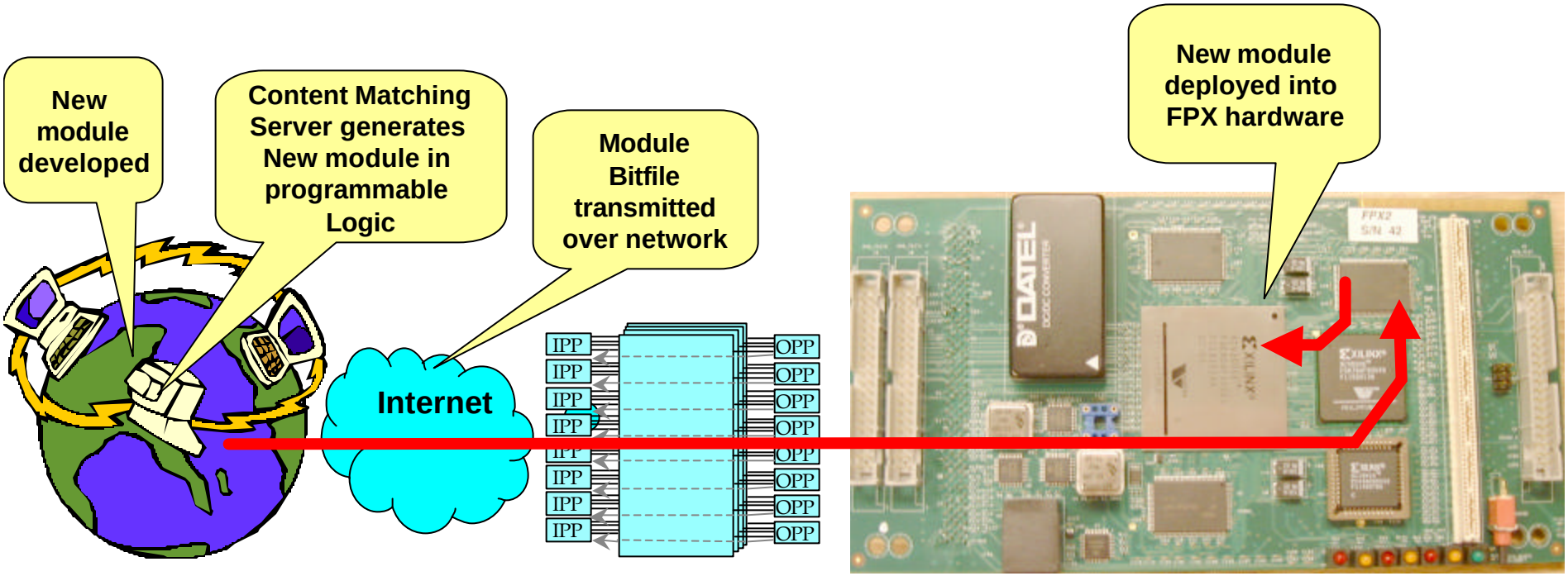
Washington University in St. Louis

Field-programmable Port Extender (FPX)



Washington University in St. Louis

Remotely reprogramming hardware over the network

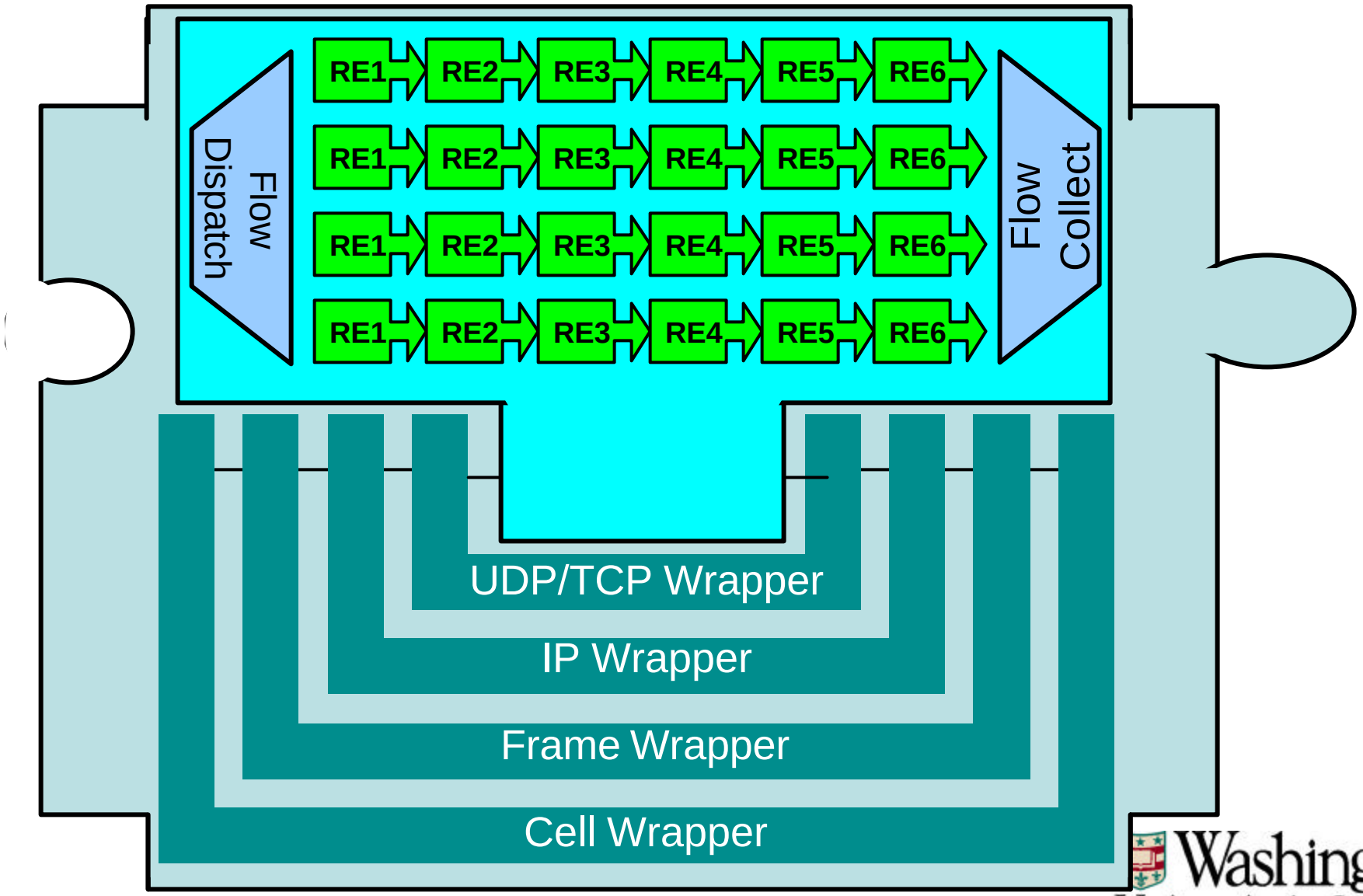


Washington University in St. Louis

Data Scanning Technologies

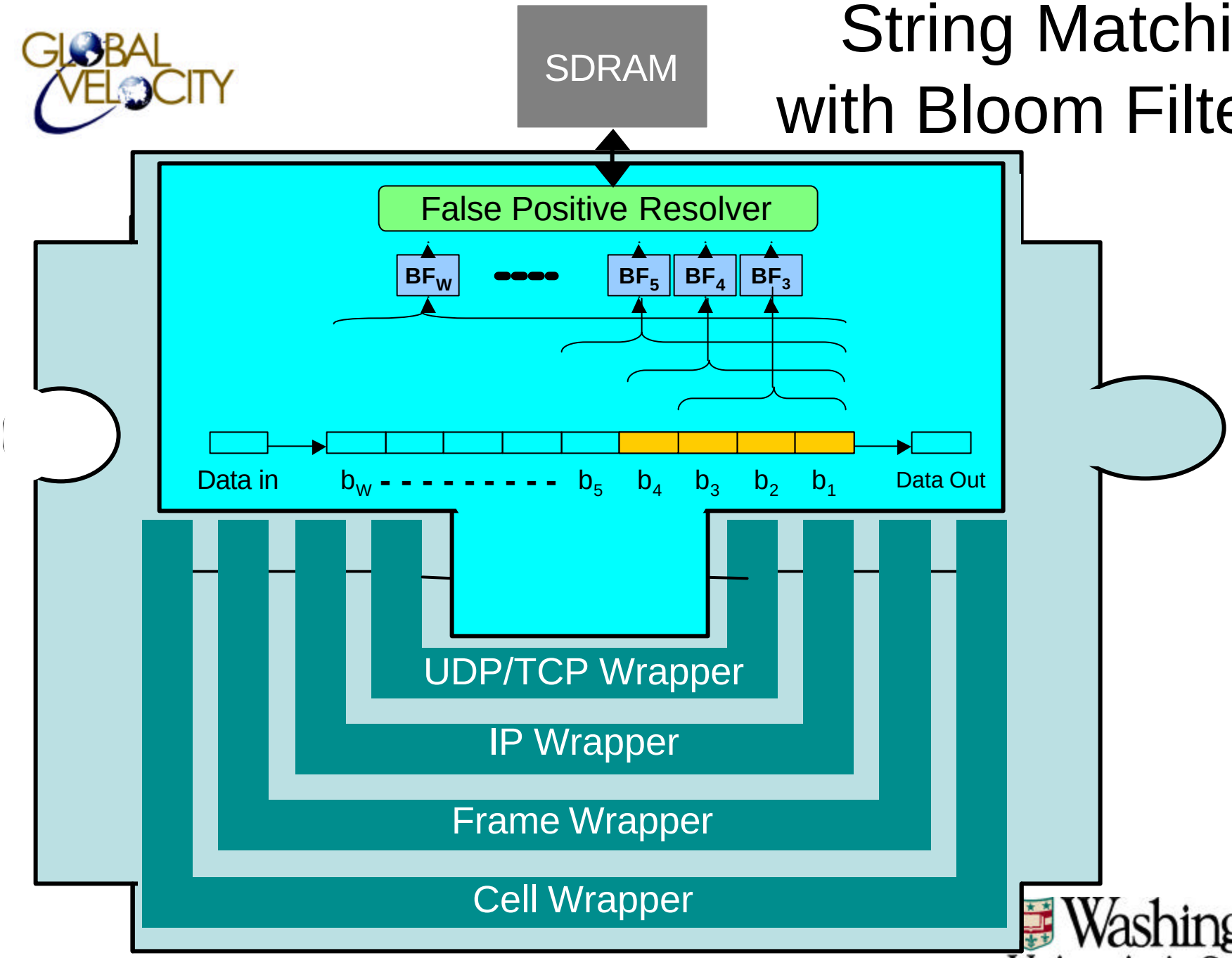
- Protocol Processing
  - Layered Protocol Wrappers
  - Process Cells/frames/packets/flows in hardware
- Regular Expression Matching
  - Deterministic Finite Automata (DFA)
  - Dynamically programmed into FPGA logic
- Fixed String Matching
  - Bloom Filters
  - Dynamically programmed into BlockRAMs

Regular Expression Matching with Finite Automata



Moscola et al.

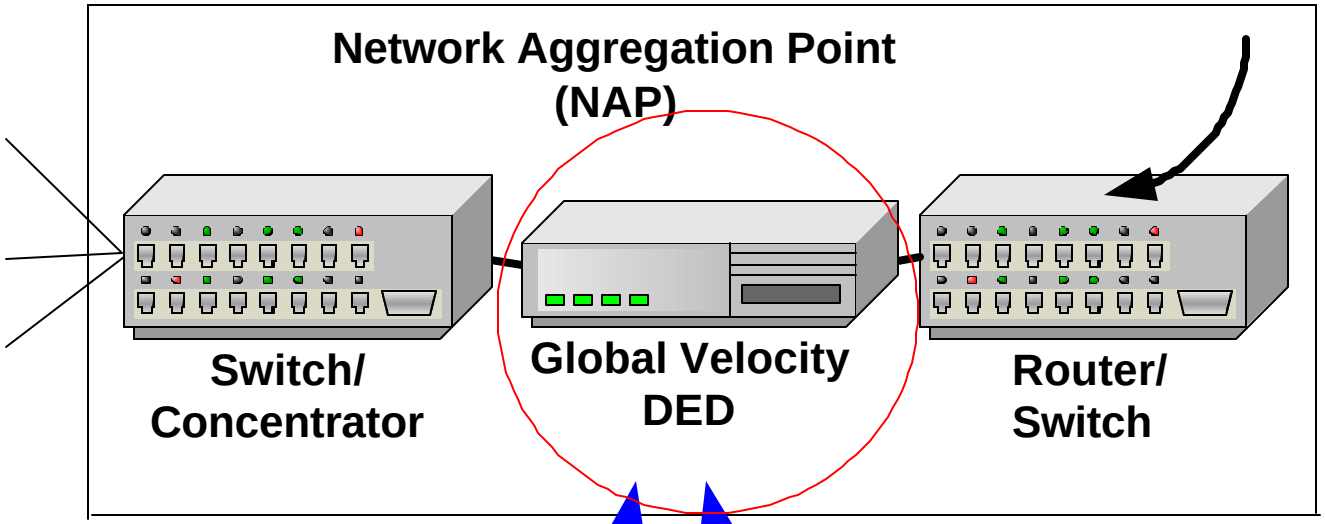
String Matching with Bloom Filters

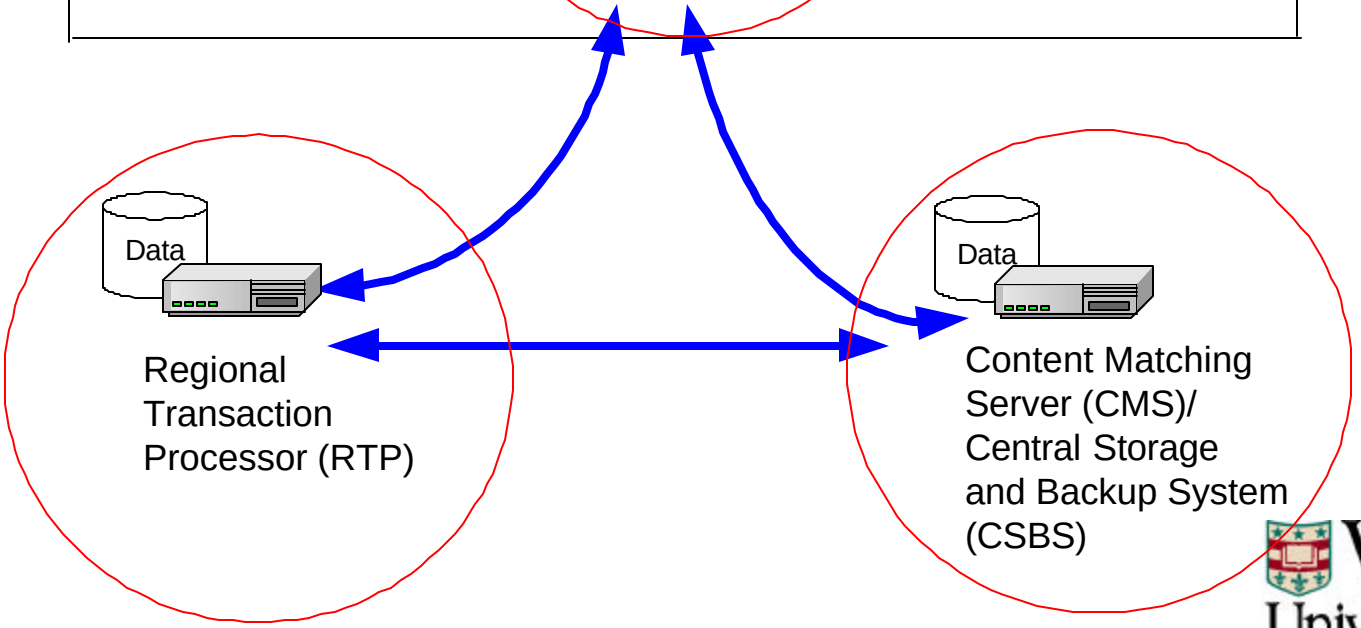


Dharmapulikar et al.

Washington University in St. Louis

Complete Protection System



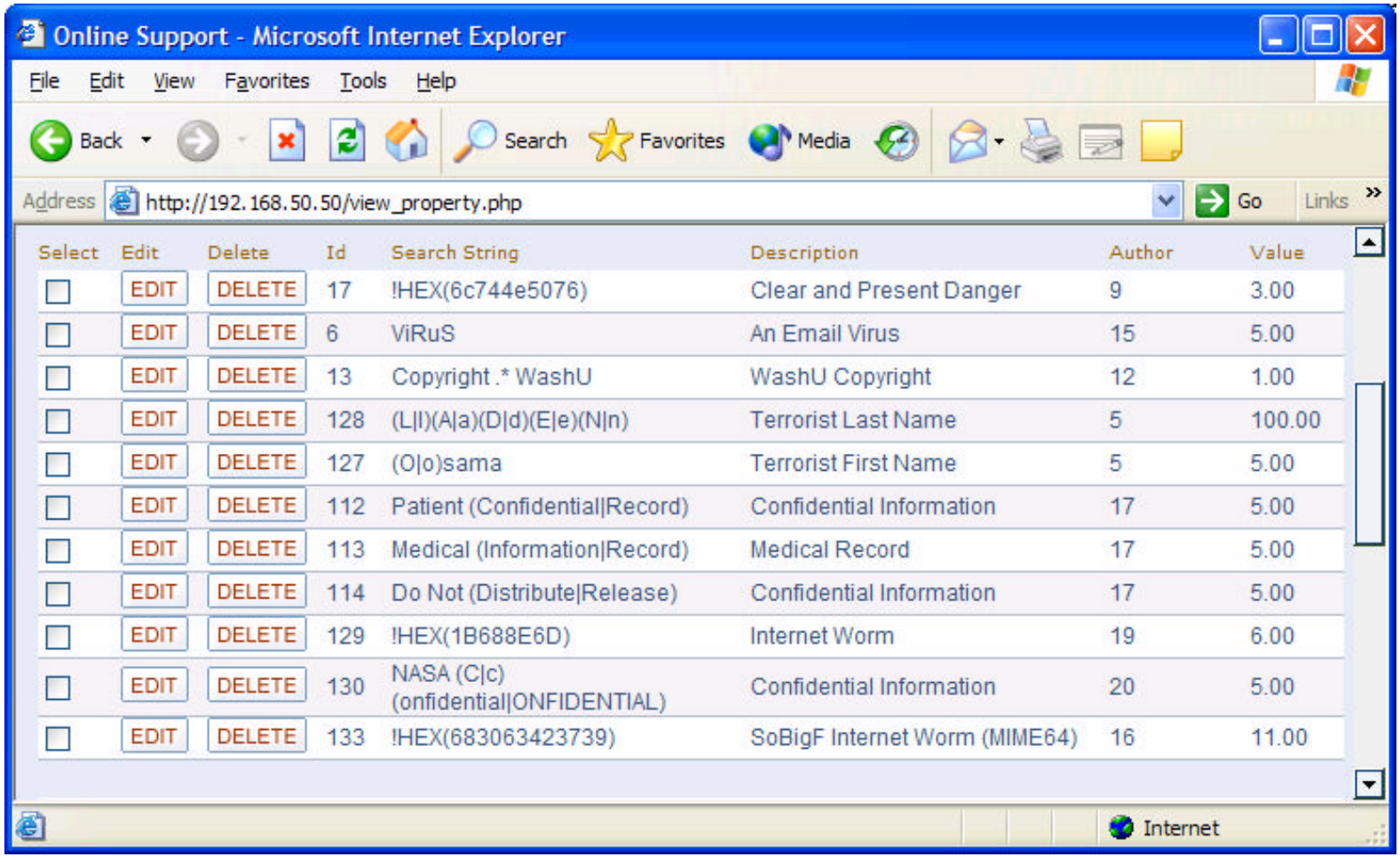


Washington University in St. Louis

System Components

- Hardware-based Data Processing
  - FPGA bitfile transferred over network to reconfigurable hardware
  - Content scanned in hardware with parallel Finite State Machines (FSMs)
  - Control messages sent over network allow blocking/unblocking of data
- Software-based System Generation
  - Web-based control and configuration
  - SQL Database stores signature patterns
  - Finite State Machines created with JLEX
  - VHDL-specified circuits generated, Instantiated, and integrated with Internet protocol processing wrappers

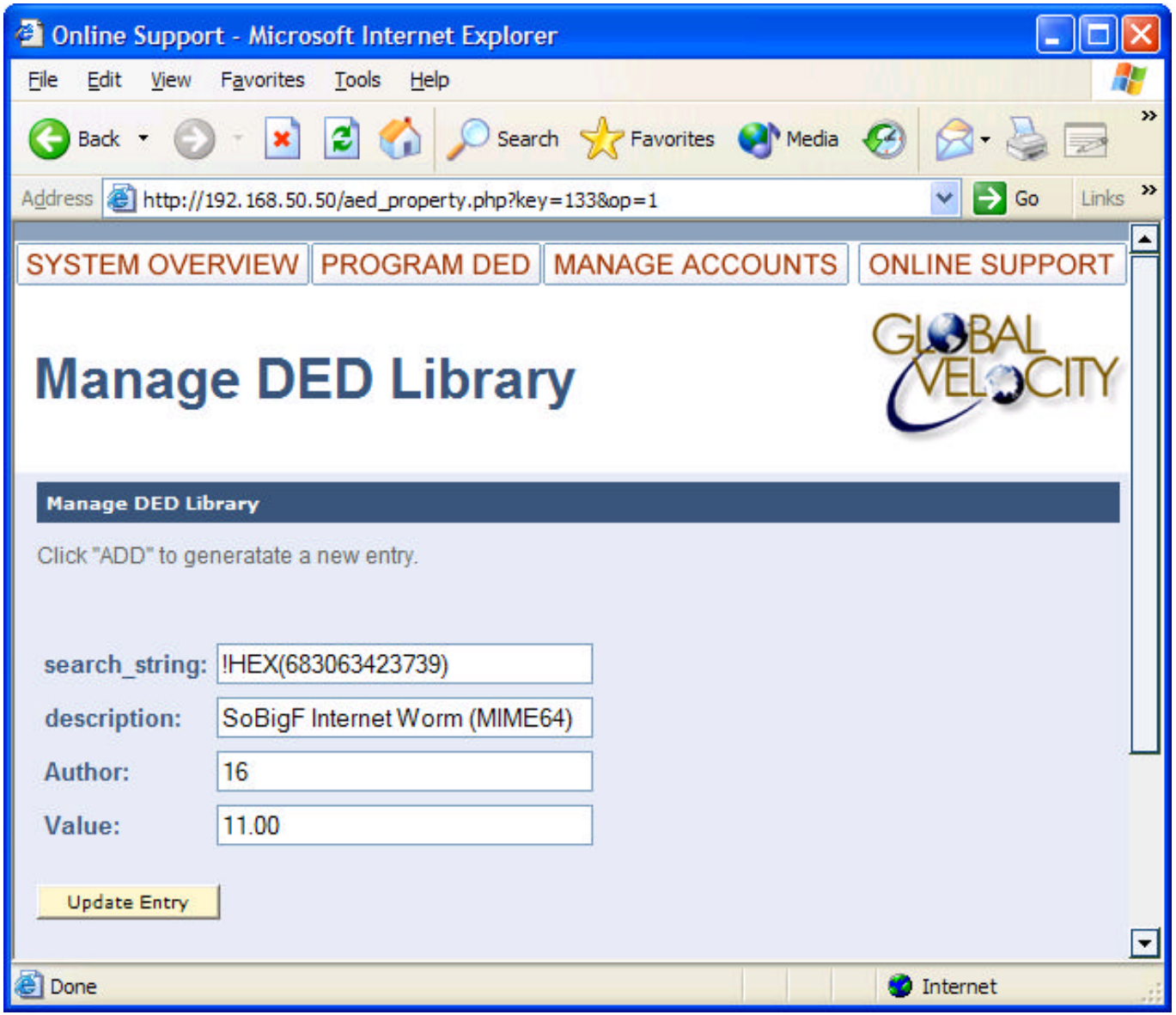
Selecting the Search Strings



Select	Edit	Delete	Id	Search String	Description	Author	Value
☐	EDIT	DELETE	17	!HEX(6c744e5076)	Clear and Present Danger	9	3.00
☐	EDIT	DELETE	6	ViRuS	An Email Virus	15	5.00
☐	EDIT	DELETE	13	Copyright .* WashU	WashU Copyright	12	1.00
☐	EDIT	DELETE	128	(L|l)(A|a)(D|d)(E|e)(N|n)	Terrorist Last Name	5	100.00
☐	EDIT	DELETE	127	(O|o)sama	Terrorist First Name	5	5.00
☐	EDIT	DELETE	112	Patient (Confidential|Record)	Confidential Information	17	5.00
☐	EDIT	DELETE	113	Medical (Information|Record)	Medical Record	17	5.00
☐	EDIT	DELETE	114	Do Not (Distribute|Release)	Confidential Information	17	5.00
☐	EDIT	DELETE	129	!HEX(1B688E6D)	Internet Worm	19	6.00
☐	EDIT	DELETE	130	NASA (C|c) (onfidential|ONFIDENTIAL)	Confidential Information	20	5.00
☐	EDIT	DELETE	133	!HEX(683063423739)	SoBigF Internet Worm (MIME64)	16	11.00

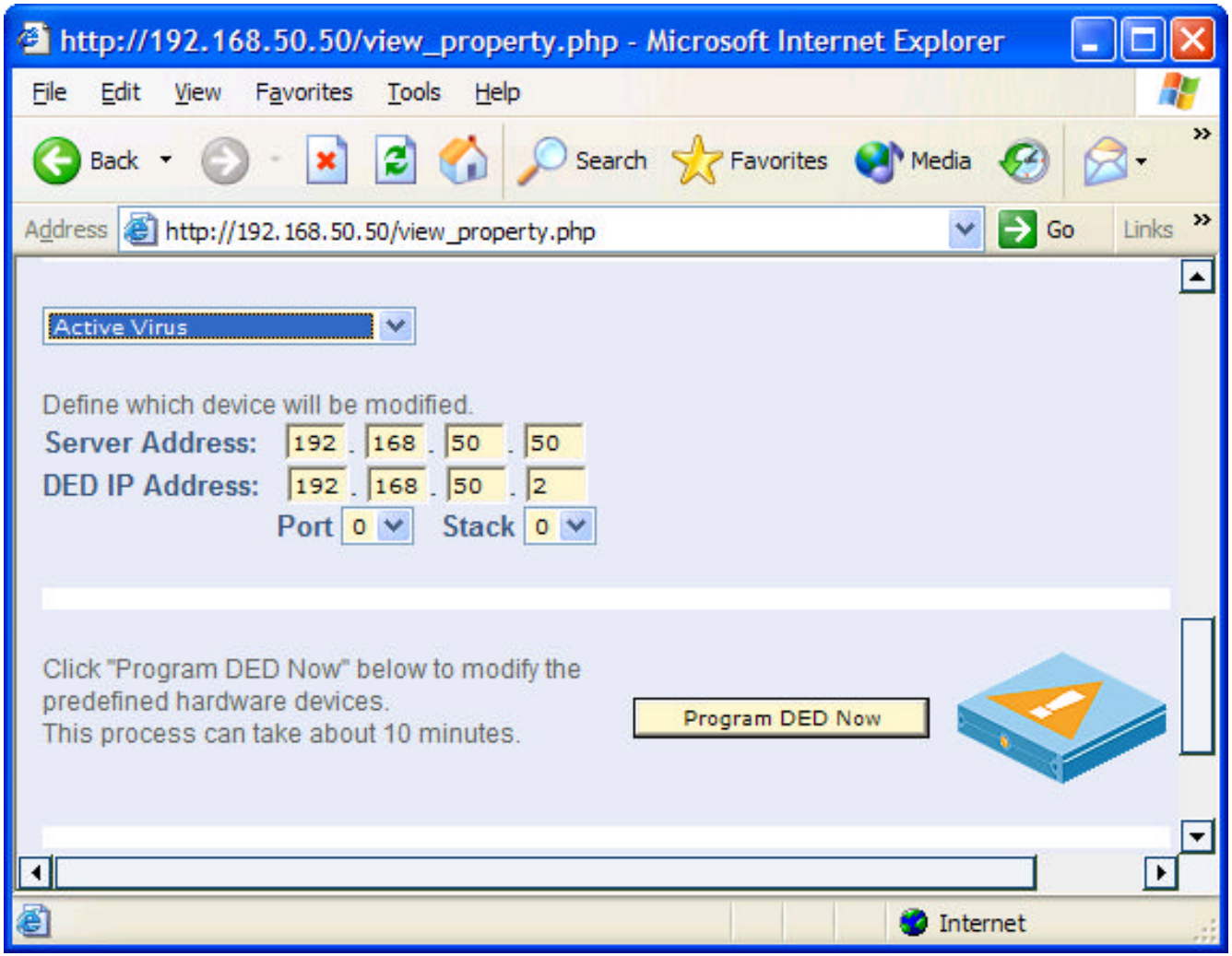
Washington University in St. Louis

Edit Search strings



Washington University in St. Louis

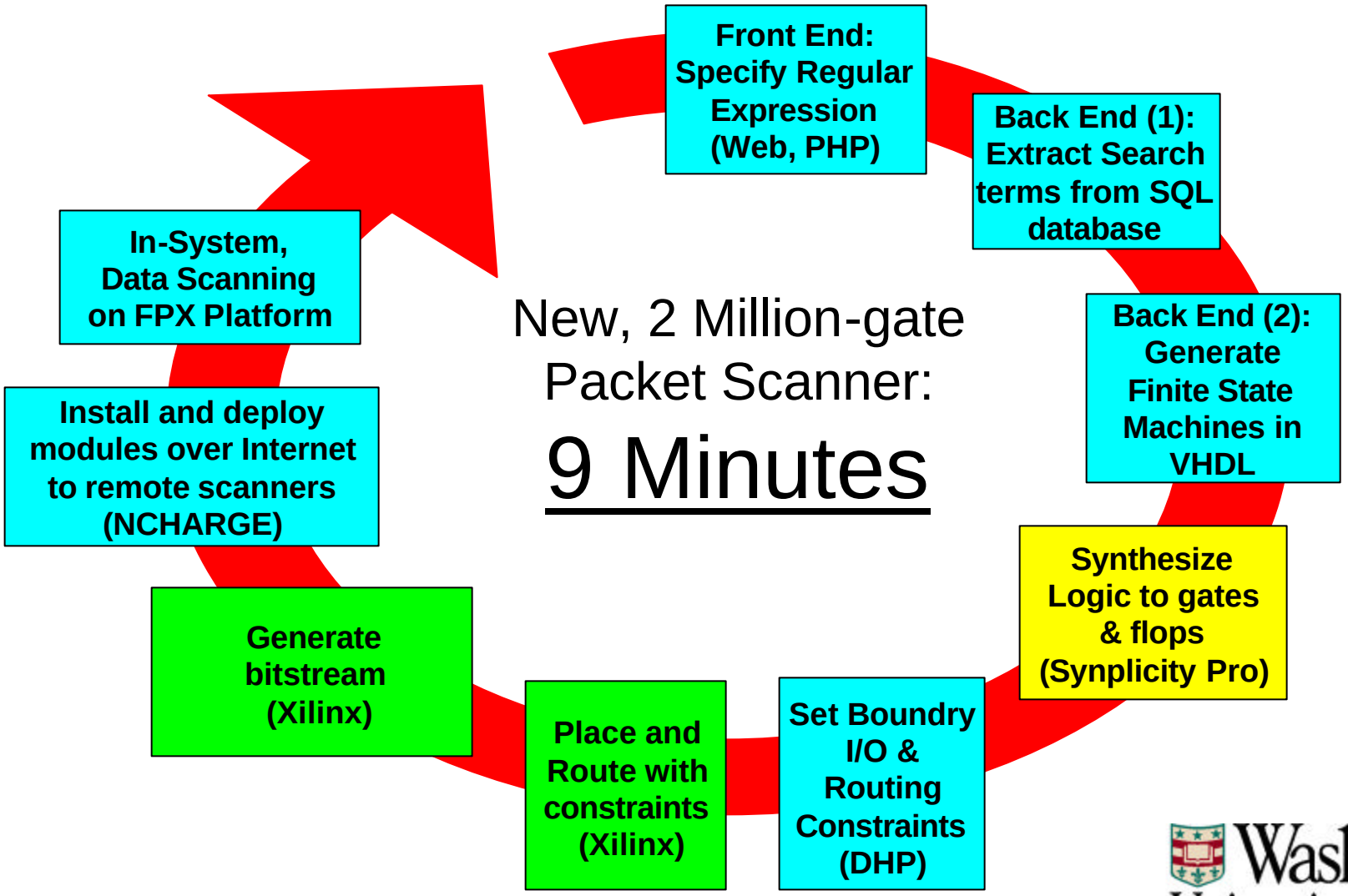
Program the Hardware



Washington University in St. Louis

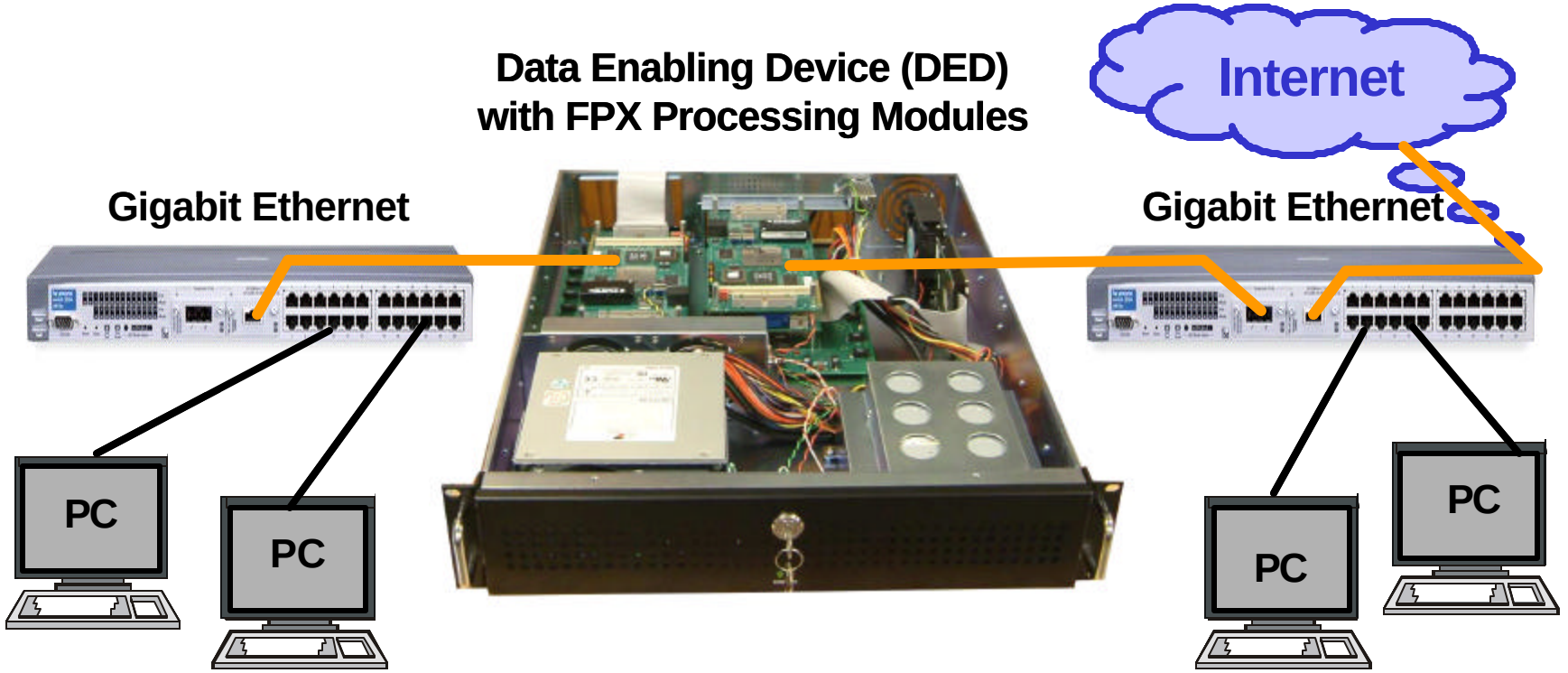
Modular Design Flow

(our contribution)



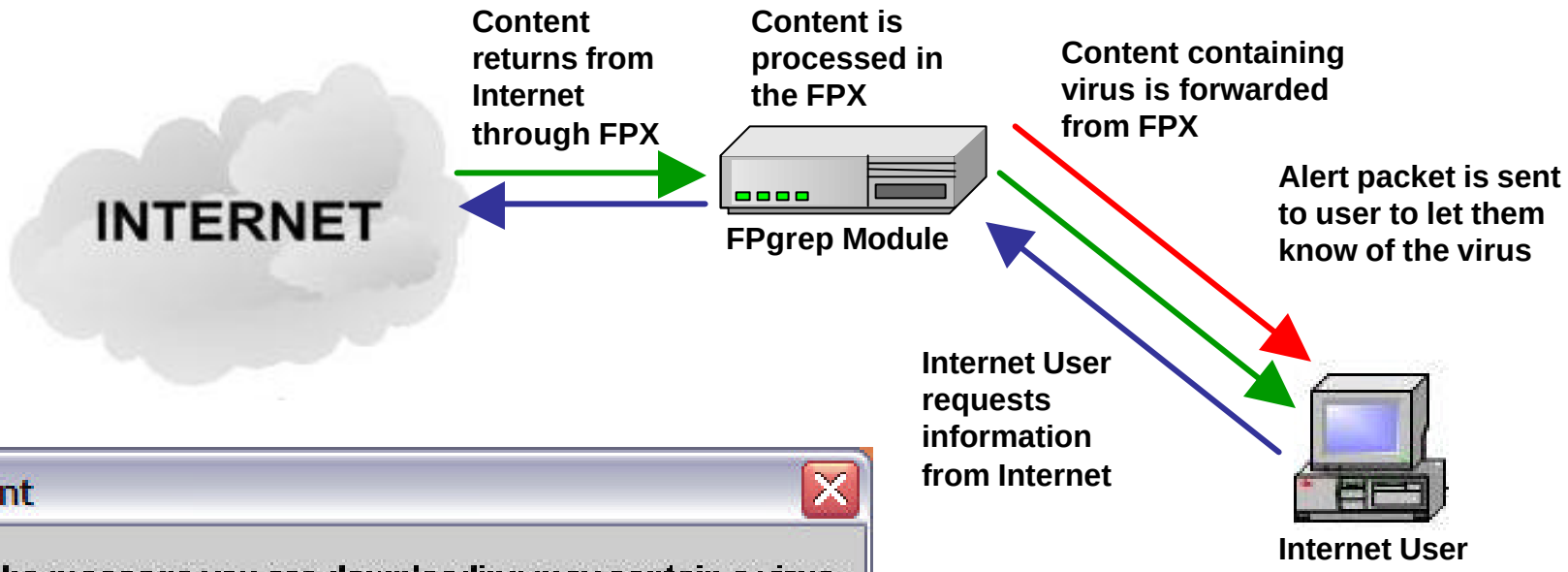
Washington University in St. Louis

Network Configuration with Gigabit Ethernet

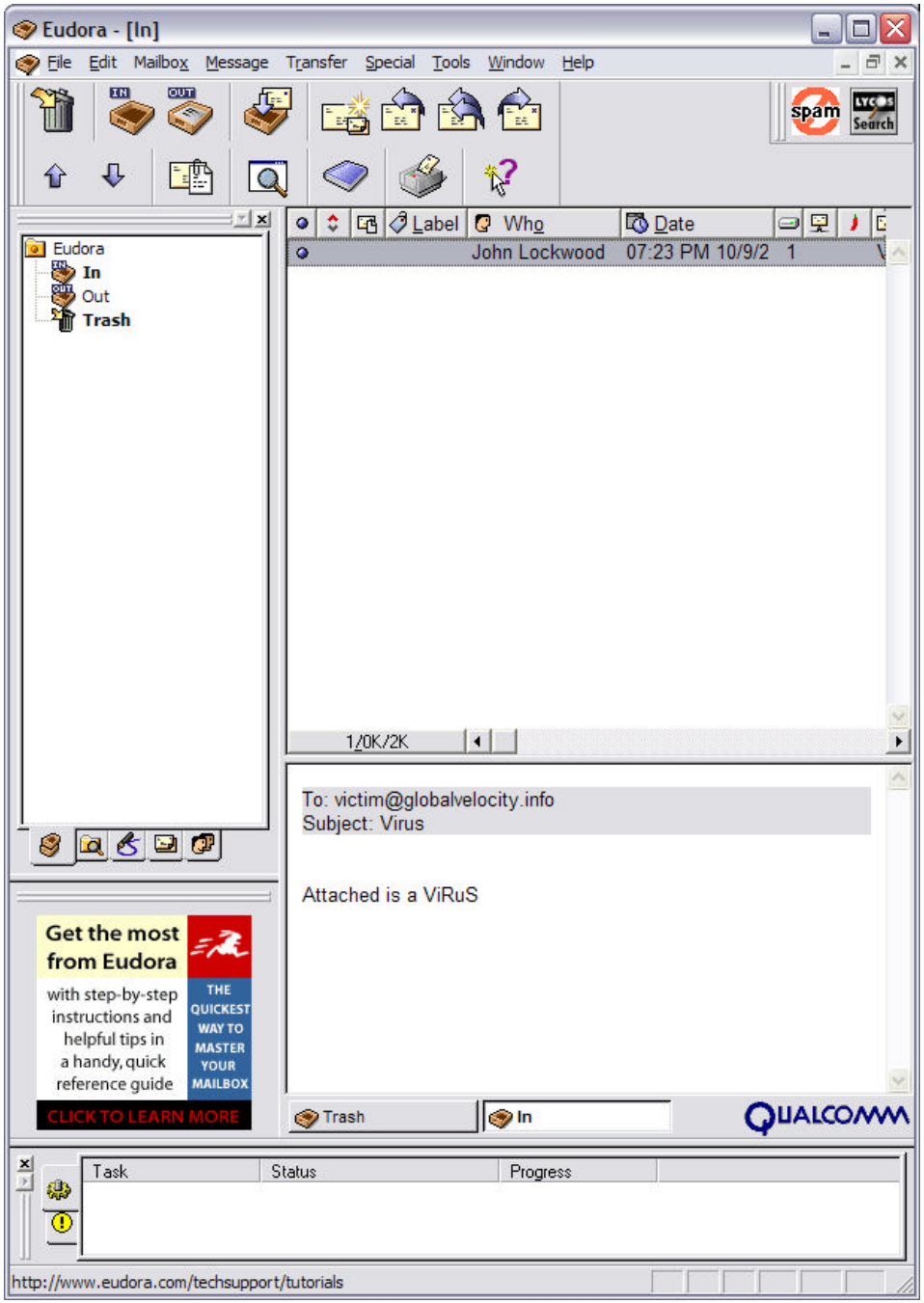


Washington University in St. Louis

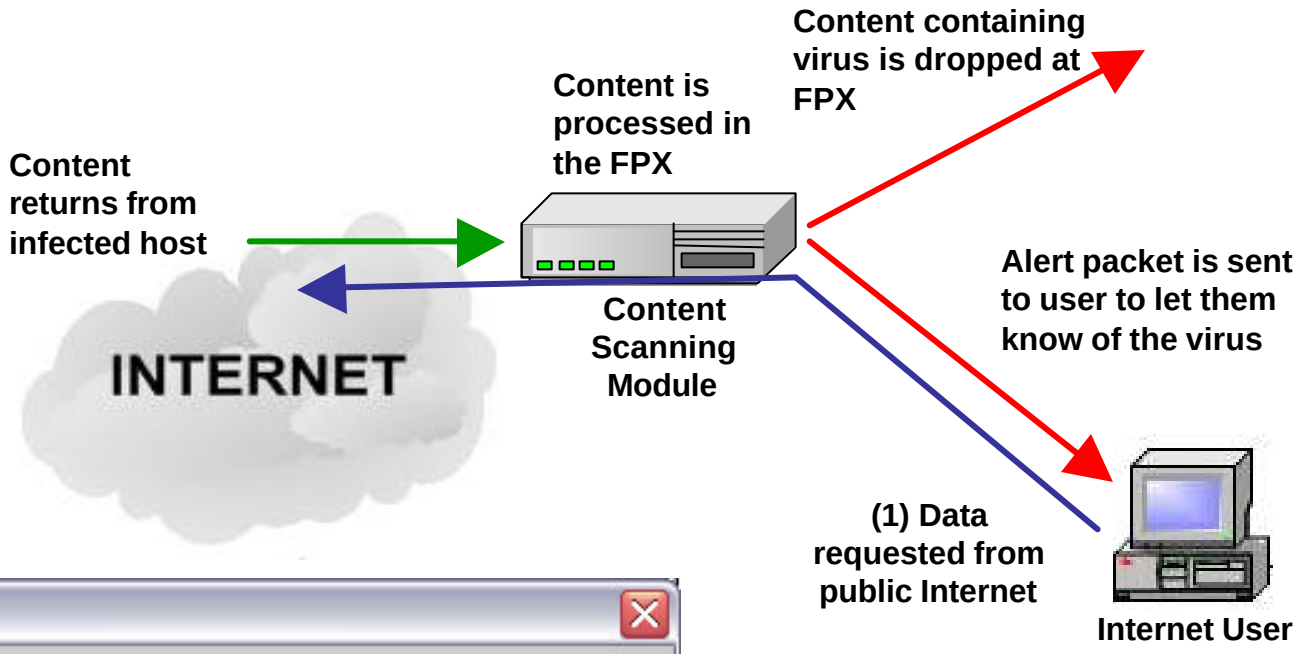
Passive Virus Protection



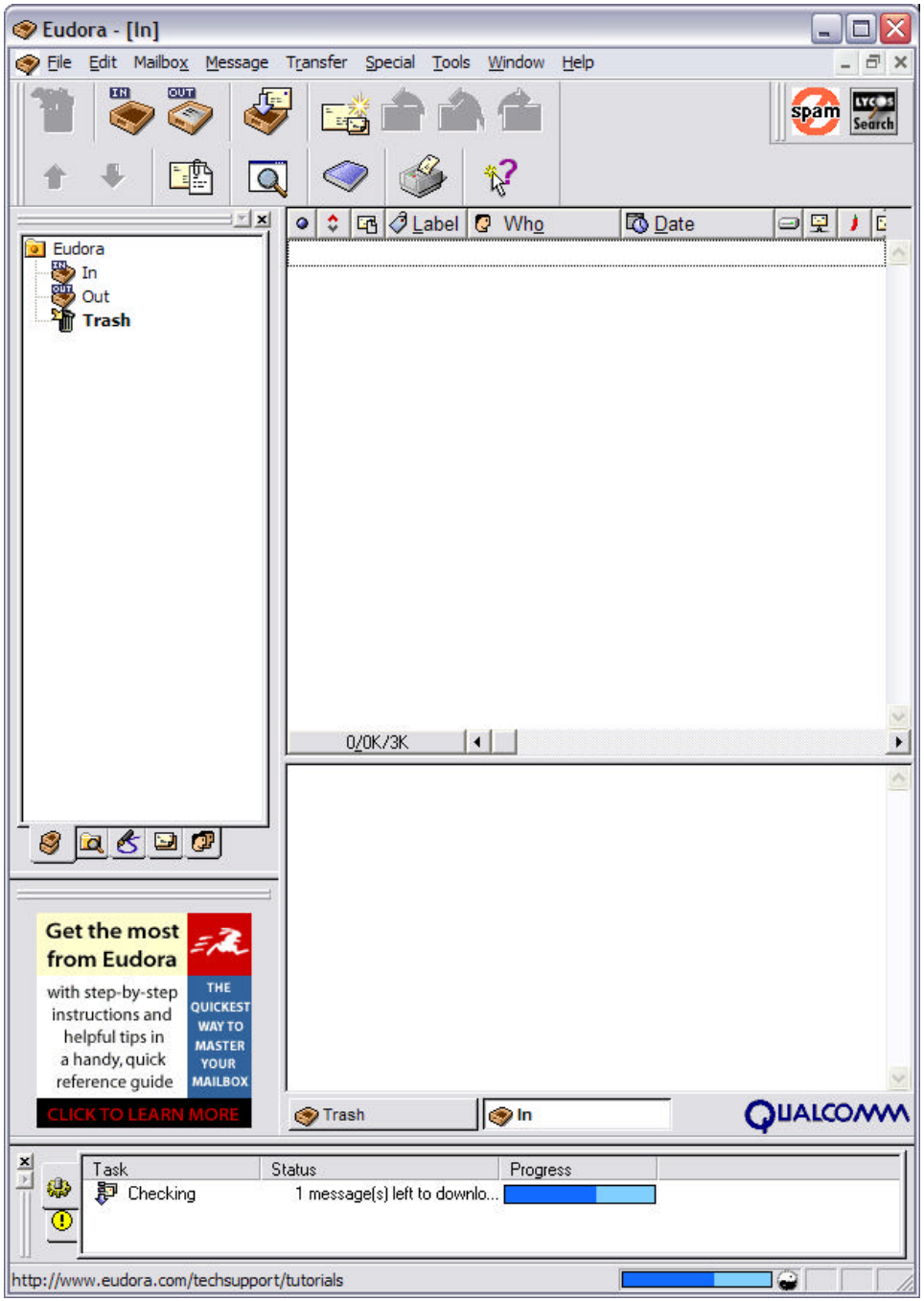
Passive Virus Example



Active Virus Protection



Active Virus Example



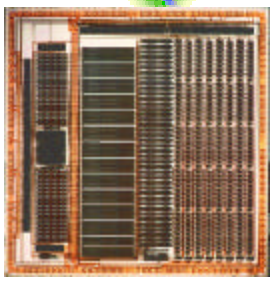
Washington University in St. Louis

Other Applications

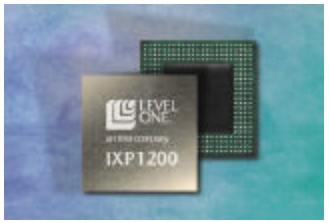
- Prevent unauthorized release of data
  - Secure Classified documents
  - Lock medical documents for Health Insurance Portability and Accountability Act (HIPAA)
- Avoid liability for misuse of network
  - Copyright infringement
  - Pornography in the workplace

Content Scanning Technologies

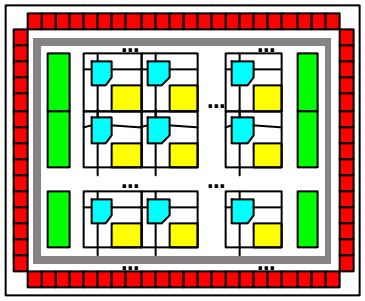
- General Purpose Microprocessors
  - ✍ Fully Reprogrammable
  - ✘ Sequential Processing



- Custom Packet Processing Hardware
  - Highly concurrent processing
  - ✘ Static Functionality



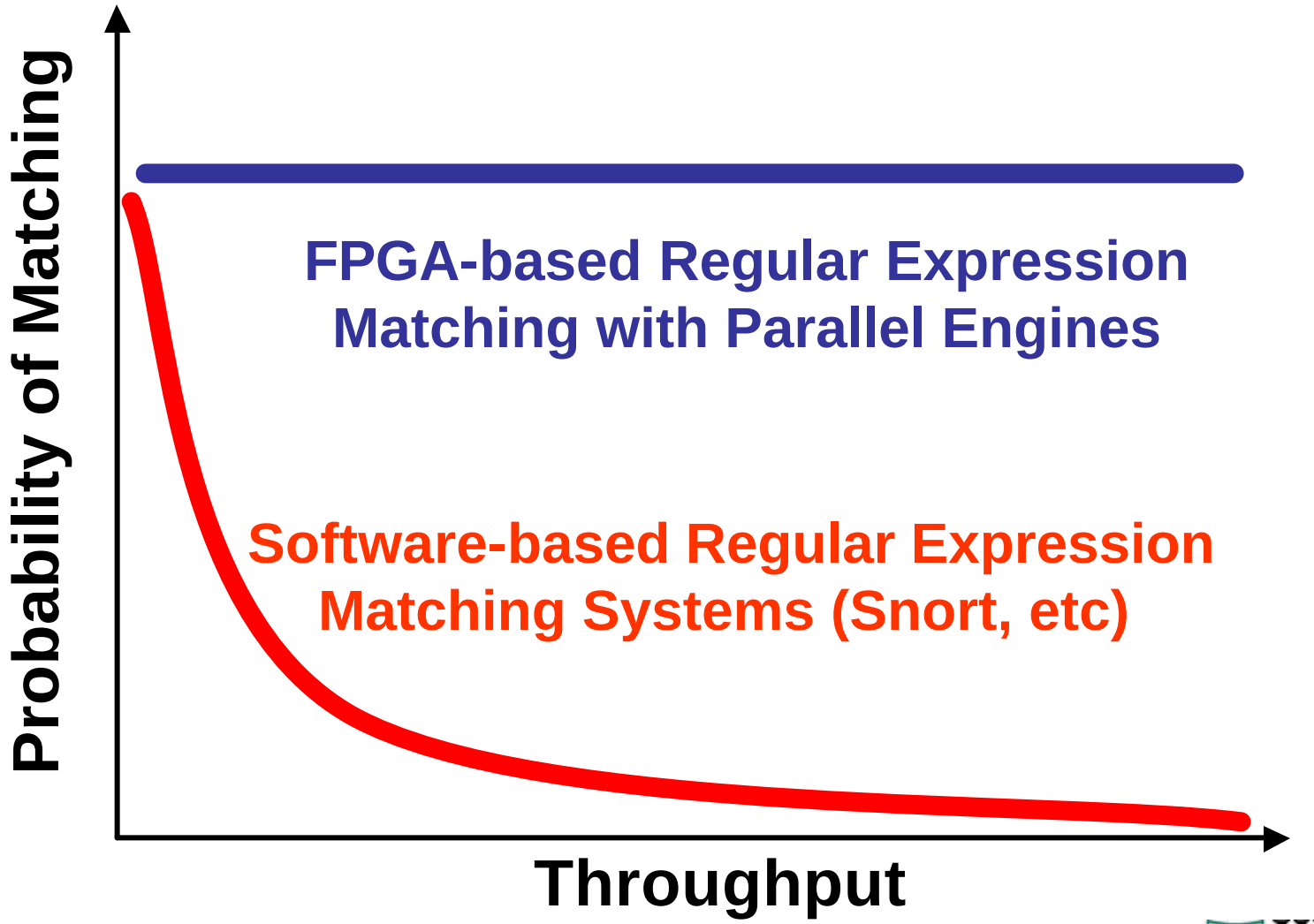
- Network Processors
  - ✍ Mostly Reprogrammable
  - Some concurrent processing (8-32 cores)



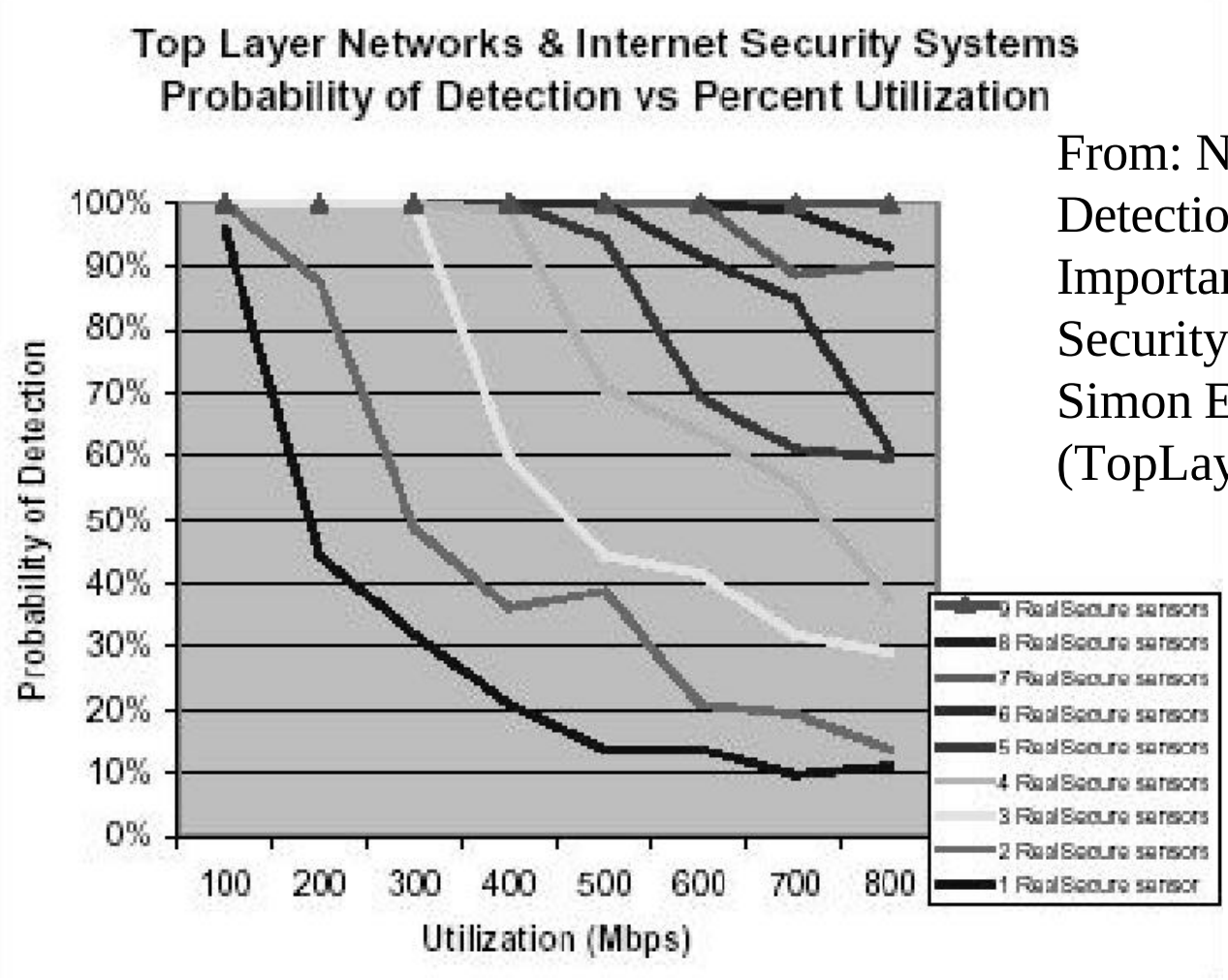
- Reconfigurable Hardware
  - ✍ Fully Programmable
  - ✍ Highly concurrent processing

Washington University in St. Louis

Performance



Actual Software Performance



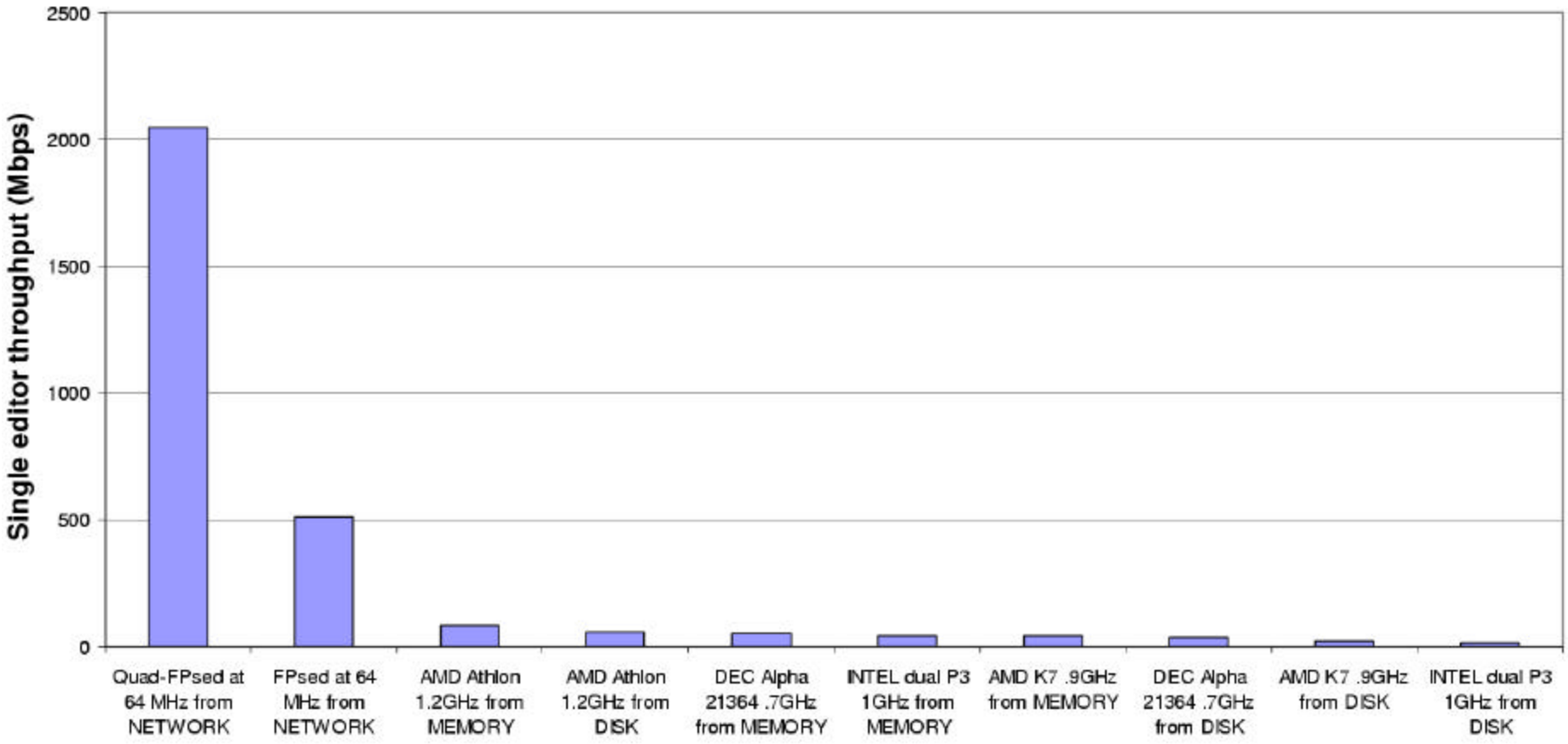
From: Network Intrusion Detection Systems: Important IDS Network Security Vulnerabilities by Simon Edwards (TopLayer.com)

Washington University in St. Louis

Throughput Comparison

- Sed was run on different Linux PCs
  - Dual Intel Pentium III @ 1 GHz
    - 13.7 Mbps when data is read from disk
    - 32.72 Mbps when data is read from memory
  - Alpha 21364 @ 667 MHz
    - 36 Mbps when data is read from disk
    - 50.4 Mbps when data is read from memory
- Software results are 40x slower than FPsed

String Processing Benchmarks (measured results for SED)



Washington University in St. Louis

Results

- Content Scanning Platform Implemented
  - Scans Internet packets for virus or Internet worm signatures using reconfigurable hardware
  - Generates prompts when matching content is found
- Content Matching Server Implemented
  - Automatically generates FPGA from regular expressions selected from database
- Regional Transaction Processor implemented
  - Tracks propagation of Internet worms and viruses
- Reduces the spread of malware from months to minutes

Acknowledgements

- Washington University
  - Faculty
    - John Lockwood
    - Ronald Loui
    - Jon Turner
  - Graduate Students
    - Mike Attig
    - Sarang Dharmapurikar
    - David Lim
    - Jing Lu
    - Bharath Madhusudan
    - James Moscola
    - Chris Neely
    - David Schuehler
    - Todd Sproull
    - David Taylor
    - Haoyu Song
    - Chris Zuver
- Industry Research Partners
  - Matthew Kulig (Global Velocity)
  - David Reddick (Global Velocity)
  - Tim Brooks (Global Velocity)
- Government Partners
  - National Science Foundation
- Hardware Vendors
  - David Parlour (Xilinx)
- Visiting Faculty and Students
  - Edson Horta
  - Florian Braun
  - Carlos Macian